

KB #3
PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/831,846

DATE: 09/27/2001

TIME: 11:56:30

Input Set : A:\06501-081001.txt

Output Set: N:\CRF3\09272001\I831846.raw

ENTERED

4 <110> APPLICANT: Funahashi, Shin-ichi
5 Miyata, Shoji
6 Nomura, Nobuo
7 Nagase, Takahiro
8 Ohara, Osamu
11 <120> TITLE OF INVENTION: NOVEL GENE ENCODING BRAIN-SPECIFIC MEMBRANE PROTEIN
13 <130> FILE REFERENCE: 06501-081001
15 <140> CURRENT APPLICATION NUMBER: 09/831,846
C--> 16 <141> CURRENT FILING DATE: 2001-08-21
18 <150> PRIOR APPLICATION NUMBER: PCT/JP99/06449
19 <151> PRIOR FILING DATE: 1999-11-18
21 <150> PRIOR APPLICATION NUMBER: JP 10/331727
22 <151> PRIOR FILING DATE: 1998-11-20
24 <160> NUMBER OF SEQ ID NOS: 7
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 3144
28 <212> TYPE: DNA
29 <213> ORGANISM: Homo sapiens
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (466)..(2832)
35 <400> SEQUENCE: 1
36 gcctggctcc ctctcgctga gacacacata cactcacaca tacacaaccc ggcaggctcg 60
38 tctgaacttg aagacacccc acattccaag atgcccgagg ttctctgggaa tgctgggggt 120
40 tcttcgatcc ggaaaatcct accggcatcc tcctagggag ggattattat tattattttt 180
42 ctttaatctg gaagagaaga gaacaagttg tgcttttccc cccttcttct tgctaaacgc 240
44 catggatata actgaataag cggctcaggg ctttccccgc gtggacgtcc gaggccacca 300
46 tctgcctgca ttcgccggag ccgccggagg gtttagctcg agtctgtctc gggcggggaa 360
48 ggatgcgtgg ccgagccggg gagcccgggc gccccgcgga gccggcctcg gtgccacca 420
50 gccgggggta gatgctgcct cgcccaggcg ctgagtgacc agacc atg gag acc ctg 477
51 Met Glu Thr Leu
52 1
54 ctt ggt ggc ctg cta gcg ttt ggc atg gcg ttt gcc gtg gtc gac gcc 525
55 Leu Gly Gly Leu Leu Ala Phe Gly Met Ala Phe Ala Val Val Asp Ala
56 5 10 15 20
58 tgc ccc aag tac tgt gtc tgc cag aat ctg tct gag tca ctg ggg acc 573
59 Cys Pro Lys Tyr Cys Val Cys Gln Asn Leu Ser Glu Ser Leu Gly Thr
60 25 30 35
62 ctg tgc ccc tcc aag ggg ctg ctc ttt gta ccc cct gat att gac cgg 621
63 Leu Cys Pro Ser Lys Gly Leu Leu Phe Val Pro Pro Asp Ile Asp Arg
64 40 45 50
66 cgg aca gtg gag ctg cgc ctg ggc ggc aac ttc atc atc cac atc agc 669
67 Arg Thr Val Glu Leu Arg Leu Gly Gly Asn Phe Ile Ile His Ile Ser
68 55 60 65
70 cgc cag gac ttt gcc aac atg acg ggg ctg gtg gac ctg acc ctg tcc 717
71 Arg Gln Asp Phe Ala Asn Met Thr Gly Leu Val Asp Leu Thr Leu Ser
72 70 75 80

RAW SEQUENCE LISTING

DATE: 09/27/2001

PATENT APPLICATION: US/09/831,846

TIME: 11:56:30

Input Set : A:\06501-081001.txt

Output Set: N:\CRF3\09272001\I831846.raw

```

74 agg aac acc atc agc cac atc cag ccc ttt tcc ttt ctg gac ctc gag      765
75 Arg Asn Thr Ile Ser His Ile Gln Pro Phe Ser Phe Leu Asp Leu Glu
76 85                               90                               95                               100
78 agc ctc cgc tcc ctg cat ctt gac agc aat cgg ctg cca agc ctt ggg      813
79 Ser Leu Arg Ser Leu His Leu Asp Ser Asn Arg Leu Pro Ser Leu Gly
80                               105                               110                               115
82 gag gac acc ctc cgg ggc ctg gtc aac ctg cag cac ctt atc gtg aac      861
83 Glu Asp Thr Leu Arg Gly Leu Val Asn Leu Gln His Leu Ile Val Asn
84                               120                               125                               130
86 aac aac cag ctg ggc ggc atc gca gat gag gct ttt gag gac ttc ctg      909
87 Asn Asn Gln Leu Gly Gly Ile Ala Asp Glu Ala Phe Glu Asp Phe Leu
88                               135                               140                               145
90 ctg aca ttg gag gat ctg gac ctc tcc tac aac aac ctc cat ggc ctg      957
91 Leu Thr Leu Glu Asp Leu Asp Leu Ser Tyr Asn Asn Leu His Gly Leu
92                               150                               155                               160
94 ccg tgg gac tcc gtg cga cgc atg gtc aac ctc cac cag ctg agc ctg      1005
95 Pro Trp Asp Ser Val Arg Arg Met Val Asn Leu His Gln Leu Ser Leu
96 165                               170                               175                               180
98 gac cac aac ctg ctg gat cac atc gcc gag ggc acc ttt gca gac ctg      1053
99 Asp His Asn Leu Leu Asp His Ile Ala Glu Gly Thr Phe Ala Asp Leu
100                               185                               190                               195
102 cag aaa ctg gcc cgc ctg gat ctc acc tcc aat cgg ctg cag aag ctg      1101
103 Gln Lys Leu Ala Arg Leu Asp Leu Thr Ser Asn Arg Leu Gln Lys Leu
104                               200                               205                               210
106 ccc cct gat ccc atc ttt gcc cgc tcc cag gct tcg gct ttg aca gcc      1149
107 Pro Pro Asp Pro Ile Phe Ala Arg Ser Gln Ala Ser Ala Leu Thr Ala
108                               215                               220                               225
110 aca ccc ttt gcc cca ccc ttg tcc ttt agt ttt ggg ggt aac cca ctt      1197
111 Thr Pro Phe Ala Pro Pro Leu Ser Phe Ser Phe Gly Gly Asn Pro Leu
112                               230                               235                               240
114 cac tgc aat tgt gag ctt ctc tgg ctg cgg agg ctc gag cgg gac gat      1245
115 His Cys Asn Cys Glu Leu Leu Trp Leu Arg Arg Leu Glu Arg Asp Asp
116 245                               250                               255                               260
118 gac ctg gaa acc tgt ggc tcc cca ggg ggc ctc aag ggt cgc tac ttc      1293
119 Asp Leu Glu Thr Cys Gly Ser Pro Gly Gly Leu Lys Gly Arg Tyr Phe
120                               265                               270                               275
122 tgg cat gtg cgt gag gag gag ttt gtg tgc gag ccg cct ctc atc acc      1341
123 Trp His Val Arg Glu Glu Glu Phe Val Cys Glu Pro Pro Leu Ile Thr
124                               280                               285                               290
126 cag cac aca cac aag ttg ctg gtt ctg gag ggc cag gcg gcc aca ctc      1389
127 Gln His Thr His Lys Leu Leu Val Leu Glu Gly Gln Ala Ala Thr Leu
128                               295                               300                               305
130 aag tgc aaa gcc att ggg gac ccc agc ccc ctt atc cac tgg gta gcc      1437
131 Lys Cys Lys Ala Ile Gly Asp Pro Ser Pro Leu Ile His Trp Val Ala
132                               310                               315                               320
134 ccc gat gac cgc ctg gta ggg aac tcc tca agg acc gct gtc tat gac      1485
135 Pro Asp Asp Arg Leu Val Gly Asn Ser Ser Arg Thr Ala Val Tyr Asp
136 325                               330                               335                               340
138 aat ggc acc ctg gac atc ttc atc acc aca tct cag gac agt ggt gcc      1533

```

RAW SEQUENCE LISTING

DATE: 09/27/2001

PATENT APPLICATION: US/09/831,846

TIME: 11:56:30

Input Set : A:\06501-081001.txt

Output Set: N:\CRF3\09272001\I831846.raw

```

139 Asn Gly Thr Leu Asp Ile Phe Ile Thr Thr Ser Gln Asp Ser Gly Ala
140          345          350          355
142 ttc acc tgc att gct gcc aat gct gcc gga gag gcc acg gcc atg gtg 1581
143 Phe Thr Cys Ile Ala Ala Asn Ala Ala Gly Glu Ala Thr Ala Met Val
144          360          365          370
146 gag gtc tcc atc gtc cag ctg cca cac ctc agc aac agc acc agc cgc 1629
147 Glu Val Ser Ile Val Gln Leu Pro His Leu Ser Asn Ser Thr Ser Arg
148          375          380          385
150 act gca ccc ccc aag tcc cgc ctc tca gac atc act ggc tcc agc aag 1677
151 Thr Ala Pro Pro Lys Ser Arg Leu Ser Asp Ile Thr Gly Ser Ser Lys
152          390          395          400
154 acc agc cgg gga ggt gga ggc agt ggg ggc gga gag cct ccc aaa agc 1725
155 Thr Ser Arg Gly Gly Gly Gly Ser Gly Gly Gly Glu Pro Pro Lys Ser
156 405          410          415          420
158 ccc ccg gaa cgg gct gtg ctt gtg tct gaa gtg acc acc acc tcg gcc 1773
159 Pro Pro Glu Arg Ala Val Leu Val Ser Glu Val Thr Thr Thr Ser Ala
160          425          430          435
162 ctg gtc aag tgg tct gtc agc aag tca gca ccc cgg gtg aag atg tac 1821
163 Leu Val Lys Trp Ser Val Ser Lys Ser Ala Pro Arg Val Lys Met Tyr
164          440          445          450
166 cag ctg cag tac aac tgc tct gac gat gag gta ctg att tac agg atg 1869
167 Gln Leu Gln Tyr Asn Cys Ser Asp Asp Glu Val Leu Ile Tyr Arg Met
168          455          460          465
170 atc cca gcc tcc aac aag gcc ttc gtg gtc aac aac ctg gtg tca ggg 1917
171 Ile Pro Ala Ser Asn Lys Ala Phe Val Val Asn Asn Leu Val Ser Gly
172          470          475          480
174 act ggc tac gac ttg tgt gtg ctg gcc atg tgg gat gac aca gcc acg 1965
175 Thr Gly Tyr Asp Leu Cys Val Leu Ala Met Trp Asp Asp Thr Ala Thr
176 485          490          495          500
178 aca ctc acg gcc acc aac atc gtg ggc tgc gcc cag ttc ttc acc aag 2013
179 Thr Leu Thr Ala Thr Asn Ile Val Gly Cys Ala Gln Phe Phe Thr Lys
180          505          510          515
182 gct gac tac ccg cag tgc cag tcc atg cac agc cag att ctg ggc ggc 2061
183 Ala Asp Tyr Pro Gln Cys Gln Ser Met His Ser Gln Ile Leu Gly Gly
184          520          525          530
186 acc atg atc ctg gtc atc ggg ggc atc atc gtg gcc acg ctg ctg gtc 2109
187 Thr Met Ile Leu Val Ile Gly Gly Ile Ile Val Ala Thr Leu Leu Val
188          535          540          545
190 ttc atc gtc atc ctc atg gtg cgc tac aag gtc tgc aac cac gag gcc 2157
191 Phe Ile Val Ile Leu Met Val Arg Tyr Lys Val Cys Asn His Glu Ala
192          550          555          560
194 ccc agc aag atg gca gcg gcc gtg agc aat gtg tac tcg cag acc aac 2205
195 Pro Ser Lys Met Ala Ala Ala Val Ser Asn Val Tyr Ser Gln Thr Asn
196 565          570          575          580
198 ggc gcc cag cca ccg cct cca agc agc gca cca gcc ggg gcc ccg ccg 2253
199 Gly Ala Gln Pro Pro Pro Ser Ser Ala Pro Ala Gly Ala Pro Pro
200          585          590          595
202 cag ggc ccg ccg aag gtg gtg gtg cgc aac gag ctc ctg gac ttc acc 2301
203 Gln Gly Pro Pro Lys Val Val Val Arg Asn Glu Leu Leu Asp Phe Thr

```

RAW SEQUENCE LISTING

DATE: 09/27/2001

PATENT APPLICATION: US/09/831,846

TIME: 11:56:30

Input Set : A:\06501-081001.txt

Output Set: N:\CRF3\09272001\I831846.raw

```

204          600          605          610
206 gcc agc ctg gcc cgc gcc agt gac tcc tct tcc tcc agc tcc ctg ggc      2349
207 Ala Ser Leu Ala Arg Ala Ser Asp Ser Ser Ser Ser Ser Ser Leu Gly
208          615          620          625
210 agt ggg gag gct gcg ggg ctg gga cgg gcc ccc tgg agg atc cca ccc      2397
211 Ser Gly Glu Ala Ala Gly Leu Gly Arg Ala Pro Trp Arg Ile Pro Pro
212          630          635          640
214 tcc gcc ccg cgc ccc aag ccc agc ctt gac cgc ctg atg ggg gcc ttc      2445
215 Ser Ala Pro Arg Pro Lys Pro Ser Leu Asp Arg Leu Met Gly Ala Phe
216 645          650          655          660
218 gcc tcc ctg gac ctc aag agt cag aga aag gag gag ctg ctg gac tcc      2493
219 Ala Ser Leu Asp Leu Lys Ser Gln Arg Lys Glu Glu Leu Leu Asp Ser
220          665          670          675
222 agg act cca gcc ggg aga ggg gct ggg acg tcg gcc cgg ggc cac cac      2541
223 Arg Thr Pro Ala Gly Arg Gly Ala Gly Thr Ser Ala Arg Gly His His
224          680          685          690
226 tcg gac cga gag cca ctg ctg ggg ccc cct gcg gcc cgg gcc agg agc      2589
227 Ser Asp Arg Glu Pro Leu Leu Gly Pro Pro Ala Ala Arg Ala Arg Ser
228          695          700          705
230 ctg ctc ccc ttg ccg ttg gag ggc aag gcc aaa cgc agc cac tcc ttc      2637
231 Leu Leu Pro Leu Pro Leu Glu Gly Lys Ala Lys Arg Ser His Ser Phe
232          710          715          720
234 gac atg ggg gac ttt gct gct gcg gcg gcg gga ggg gtc gtg ccg ggc      2685
235 Asp Met Gly Asp Phe Ala Ala Ala Ala Ala Gly Gly Val Val Pro Gly
236 725          730          735          740
238 ggc tac agt cct cct cgg aag gtc tcg aac atc tgg acg aag cgc agc      2733
239 Gly Tyr Ser Pro Pro Arg Lys Val Ser Asn Ile Trp Thr Lys Arg Ser
240          745          750          755
242 ctc tct gtc aac ggc atg ctc ttg ccc ttt gag gag agt gac ctg gtg      2781
243 Leu Ser Val Asn Gly Met Leu Leu Pro Phe Glu Glu Ser Asp Leu Val
244          760          765          770
246 ggg gcc cgg ggg act ttt ggc agc tcc gaa tgg gtg atg gag agc acg      2829
247 Gly Ala Arg Gly Thr Phe Gly Ser Ser Glu Trp Val Met Glu Ser Thr
248          775          780          785
250 gtc taggtggggg tgggcatgct ccctttcctg tgcgcagggt gggagaaggg      2882
251 Val
254 gaaagaatct cactggcaag tgtttgtgga gtttccatgg tgatgtttac atccagggac      2942
256 agtttcgtct ccctgtcaat ggcctcgtgt cccccctac cccgcaacac ccacatcacc      3002
258 tccccaccac ccggccgggg tgtgctcagg gaatgtggac tcgctcaaat gccggactga      3062
260 gccctgagtg tttggaaagg cgagactccg cctttctaata cacaaatgta gcctacaagc      3122
262 aagcggcttt ggattgctta tg      3144
265 <210> SEQ ID NO: 2
266 <211> LENGTH: 789
267 <212> TYPE: PRT
268 <213> ORGANISM: Homo sapiens
270 <400> SEQUENCE: 2
271 Met Glu Thr Leu Leu Gly Gly Leu Leu Ala Phe Gly Met Ala Phe Ala
272 1          5          10          15
273 Val Val Asp Ala Cys Pro Lys Tyr Cys Val Cys Gln Asn Leu Ser Glu

```

RAW SEQUENCE LISTING

DATE: 09/27/2001

PATENT APPLICATION: US/09/831,846

TIME: 11:56:30

Input Set : A:\06501-081001.txt

Output Set: N:\CRF3\09272001\I831846.raw

```

274          20          25          30
275 Ser Leu Gly Thr Leu Cys Pro Ser Lys Gly Leu Leu Phe Val Pro Pro
276          35          40          45
277 Asp Ile Asp Arg Arg Thr Val Glu Leu Arg Leu Gly Gly Asn Phe Ile
278          50          55          60
279 Ile His Ile Ser Arg Gln Asp Phe Ala Asn Met Thr Gly Leu Val Asp
280 65          70          75          80
281 Leu Thr Leu Ser Arg Asn Thr Ile Ser His Ile Gln Pro Phe Ser Phe
282          85          90          95
283 Leu Asp Leu Glu Ser Leu Arg Ser Leu His Leu Asp Ser Asn Arg Leu
284          100         105         110
285 Pro Ser Leu Gly Glu Asp Thr Leu Arg Gly Leu Val Asn Leu Gln His
286          115         120         125
287 Leu Ile Val Asn Asn Asn Gln Leu Gly Gly Ile Ala Asp Glu Ala Phe
288          130         135         140
289 Glu Asp Phe Leu Leu Thr Leu Glu Asp Leu Asp Leu Ser Tyr Asn Asn
290 145          150         155         160
291 Leu His Gly Leu Pro Trp Asp Ser Val Arg Arg Met Val Asn Leu His
292          165         170         175
293 Gln Leu Ser Leu Asp His Asn Leu Leu Asp His Ile Ala Glu Gly Thr
294          180         185         190
295 Phe Ala Asp Leu Gln Lys Leu Ala Arg Leu Asp Leu Thr Ser Asn Arg
296          195         200         205
297 Leu Gln Lys Leu Pro Pro Asp Pro Ile Phe Ala Arg Ser Gln Ala Ser
298          210         215         220
299 Ala Leu Thr Ala Thr Pro Phe Ala Pro Pro Leu Ser Phe Ser Phe Gly
300 225          230         235         240
301 Gly Asn Pro Leu His Cys Asn Cys Glu Leu Leu Trp Leu Arg Arg Leu
302          245         250         255
303 Glu Arg Asp Asp Asp Leu Glu Thr Cys Gly Ser Pro Gly Gly Leu Lys
304          260         265         270
305 Gly Arg Tyr Phe Trp His Val Arg Glu Glu Glu Phe Val Cys Glu Pro
306          275         280         285
307 Pro Leu Ile Thr Gln His Thr His Lys Leu Leu Val Leu Glu Gly Gln
308          290         295         300
309 Ala Ala Thr Leu Lys Cys Lys Ala Ile Gly Asp Pro Ser Pro Leu Ile
310 305          310         315         320
311 His Trp Val Ala Pro Asp Asp Arg Leu Val Gly Asn Ser Ser Arg Thr
312          325         330         335
313 Ala Val Tyr Asp Asn Gly Thr Leu Asp Ile Phe Ile Thr Thr Ser Gln
314          340         345         350
315 Asp Ser Gly Ala Phe Thr Cys Ile Ala Ala Asn Ala Ala Gly Glu Ala
316          355         360         365
317 Thr Ala Met Val Glu Val Ser Ile Val Gln Leu Pro His Leu Ser Asn
318          370         375         380
319 Ser Thr Ser Arg Thr Ala Pro Pro Lys Ser Arg Leu Ser Asp Ile Thr
320 385          390         395         400
321 Gly Ser Ser Lys Thr Ser Arg Gly Gly Gly Gly Ser Gly Gly Gly Glu
322          405         410         415

```

VERIFICATION SUMMARY

DATE: 09/27/2001

PATENT APPLICATION: US/09/831,846

TIME: 11:56:31

Input Set : A:\06501-081001.txt

Output Set: N:\CRF3\09272001\I831846.raw

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date